

1. 딥러닝이란 무엇인가?



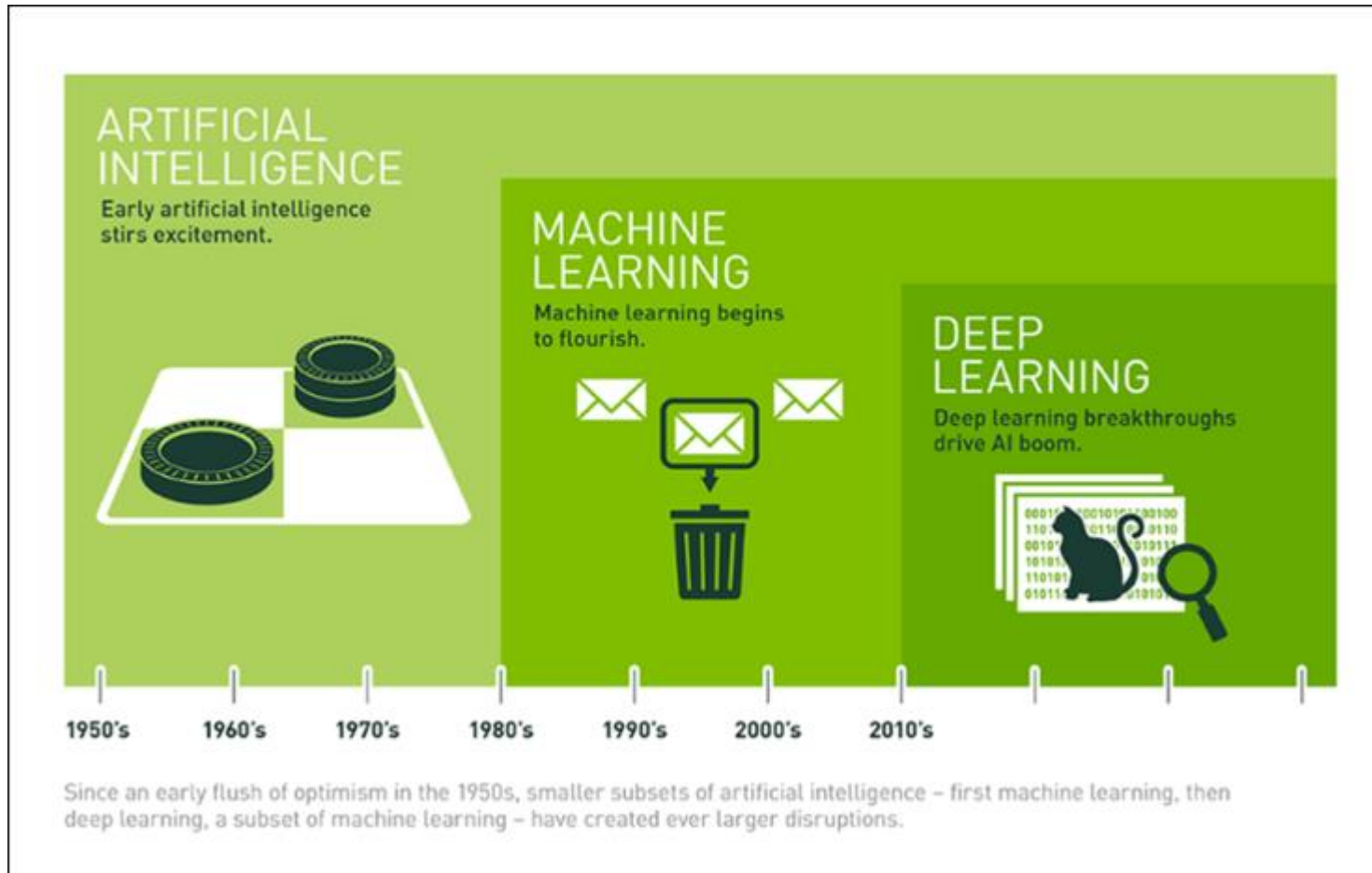
- ◆ 인공지능과 머신러닝, 딥러닝의 기본 개념을 이해한다.
- ◆ 딥러닝 이전 : 머신 러닝의 간략한 역사를 이해한다.
- ◆ 딥러닝을 왜 학습해야 하는지 이해한다.

1. 딥러닝이란 무엇인가?

1.1 인공지능과 머신러닝, 딥러닝

1.2 딥러닝 이전 : 머신 러닝의 간략한 역사

1.3 왜 딥러닝일까? 왜 지금일까?



"보통의 사람이 수행하는 지능적인 작업을 자동화하기 위한 연구 활동 - 본문"

인공지능(人工知能, 영어: artificial intelligence, AI)은 기계로부터 만들어진 지능을 말한다. - wikipedia

명시적인 규칙을 충분하게 만들어 기계의 수준을 높이는 접근 방법을 symbolic AI 라고 하며, 1950년대부터 1980년대까지 AI의 지배적인 패러다임이었습니다.

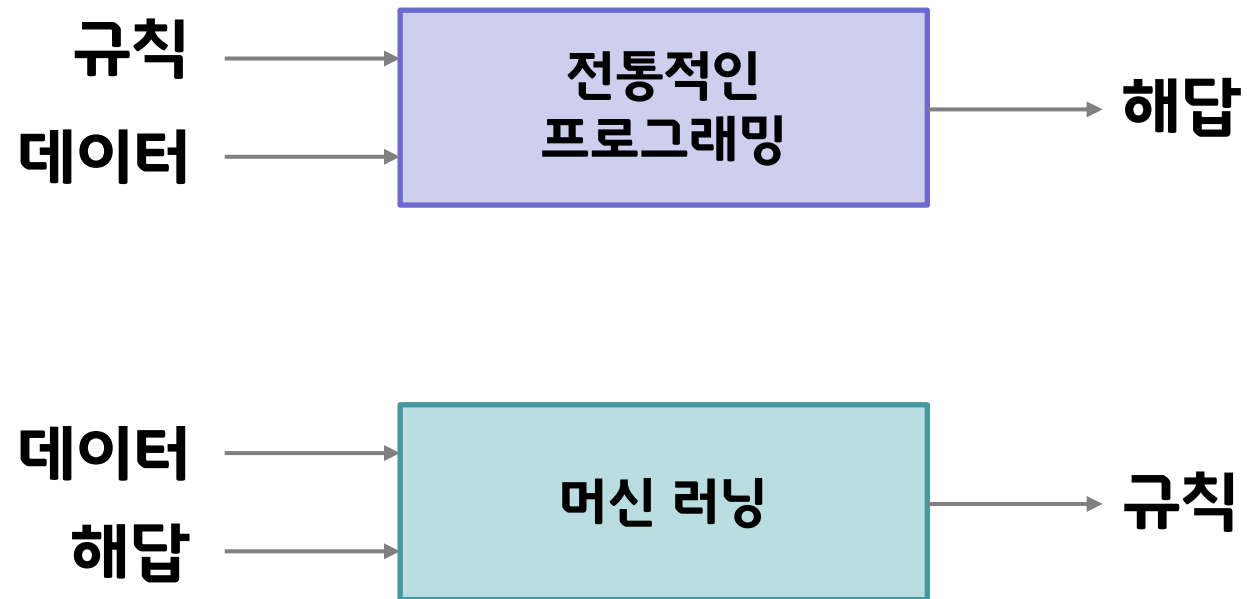
Basic concepts

- What is ML?
- What is learning?
 - supervised
 - unsupervised
- What is regression?
- What is classification?

Machine Learning

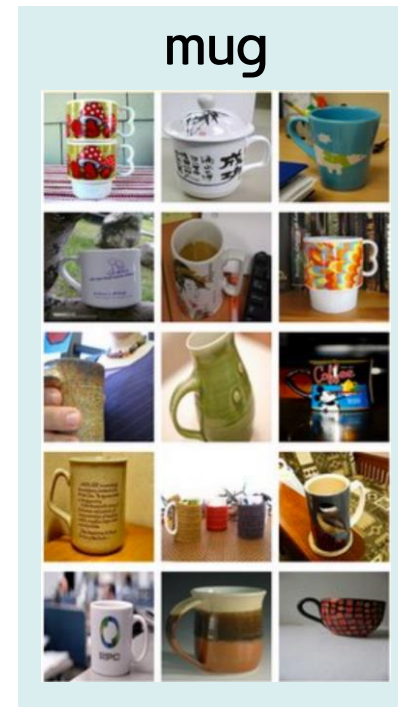
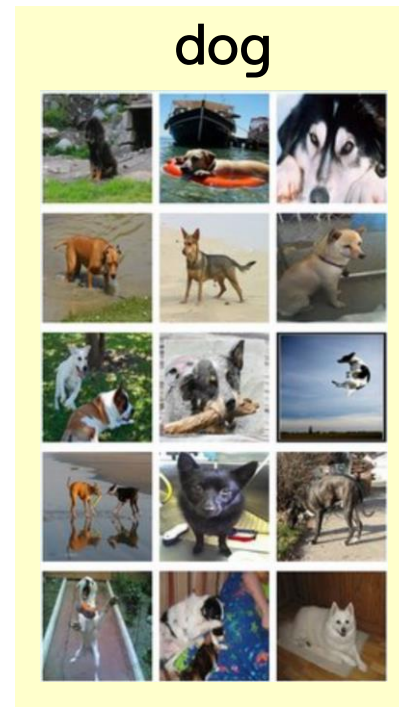
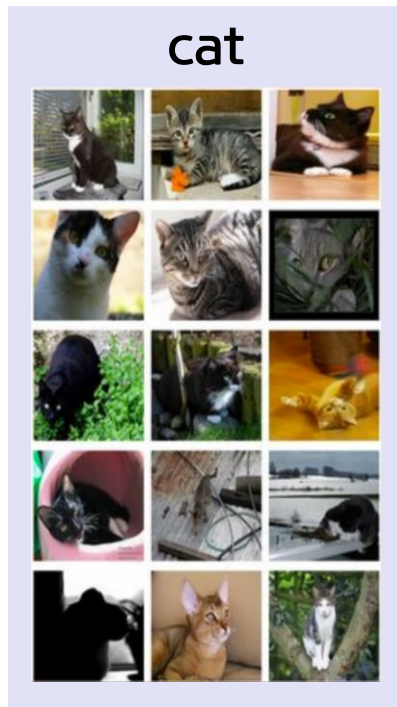
- 명시적 프로그램의 한계
 - 스팸 메일 필터 : 많은 규칙
 - 자율 주행 : 더 많은 규칙
- Machine learning :

"기계가 일일이 코드로 명시하지 않은 동작을 데이터로부터 학습하여 실행할 수 있도록 하는 알고리즘을 개발하는 연구 분야 " Arthur Samuel(1959)



Supervised/Unsupervised learning

- Supervised learning:
 - 레이블이 있는 예제로 학습



Supervised/Unsupervised learning

- Unsupervised learning:
 - 특정 입력(Input)에 대하여 올바른 정답(Right Answer)이 없는 데이터 집합이 주어지는 경우의 학습
 - 잘못된 예측에 대해 Feedback을 받고 교정할 수 없음

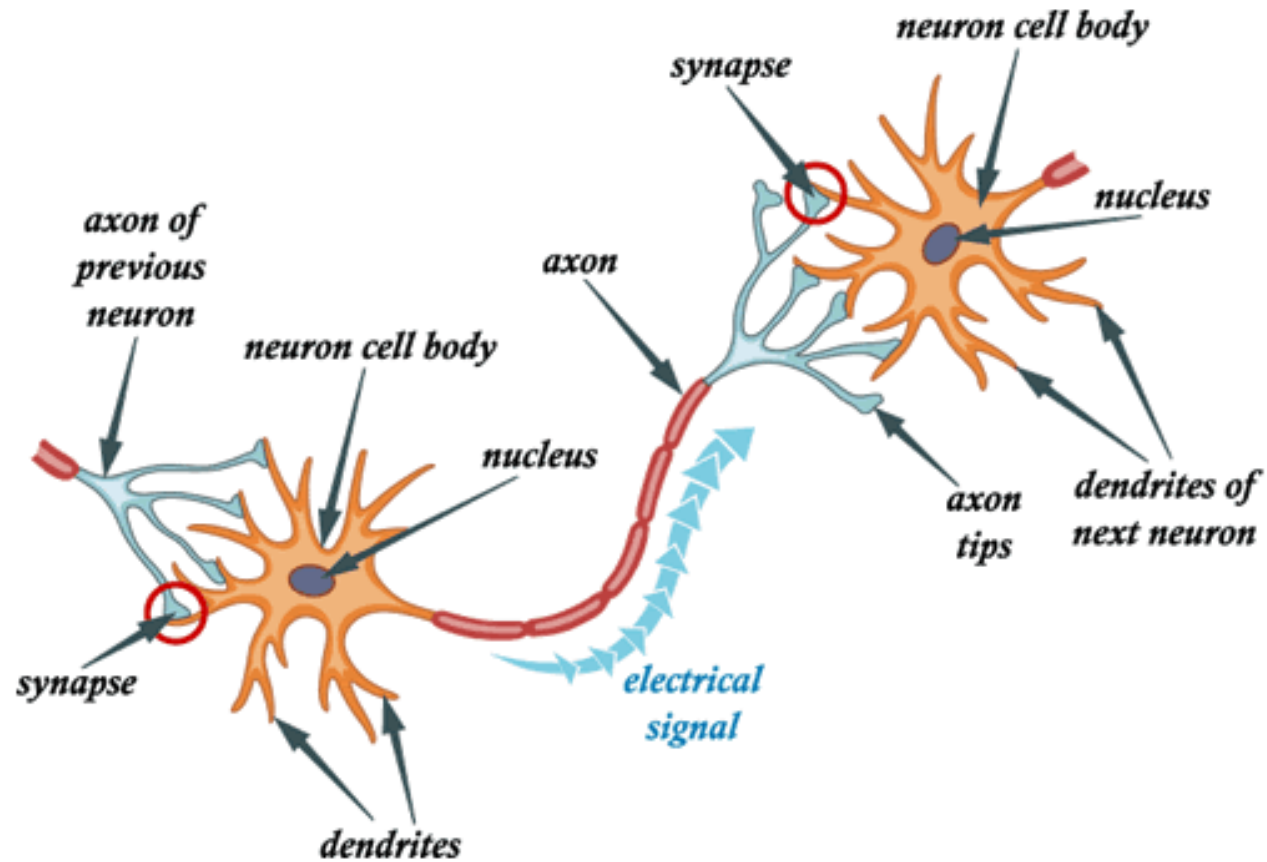
예) Google news grouping
페이스북 에서 특정 집단의 사람들을 그룹화
천체의 별 모양으로 분류

Supervised learning

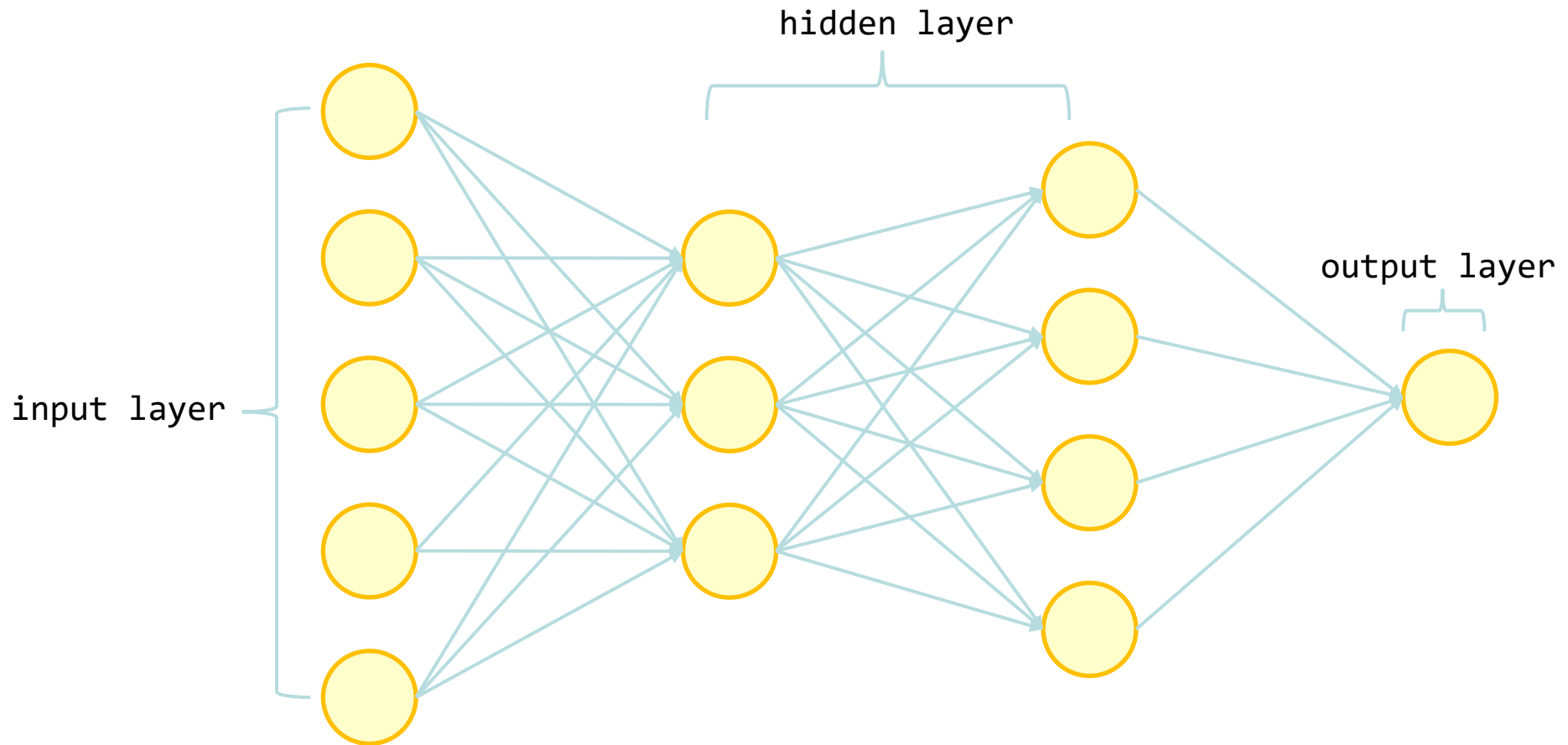
- ML에서 가장 일반적인 문제 유형
 - 시험 점수 예측 : 이전 시험에서 점수와 공부 시간
 - 이메일 스팸 필터 : 라벨이 있는 학습(스팸 또는 햄)
 - 이미지 라벨링 : 태그가 있는 이미지로 부터 학습

Types of supervised learning

- 시험 점수 예측 : 이전 시험에서 점수와 공부 시간
 - regression
- 이메일 스팸 필터 : 라벨이 있는 학습(스팸 또는 햄)
 - binary classification
- 이미지 라벨링 : 태그가 있는 이미지로 부터 학습
 - multi-label classification



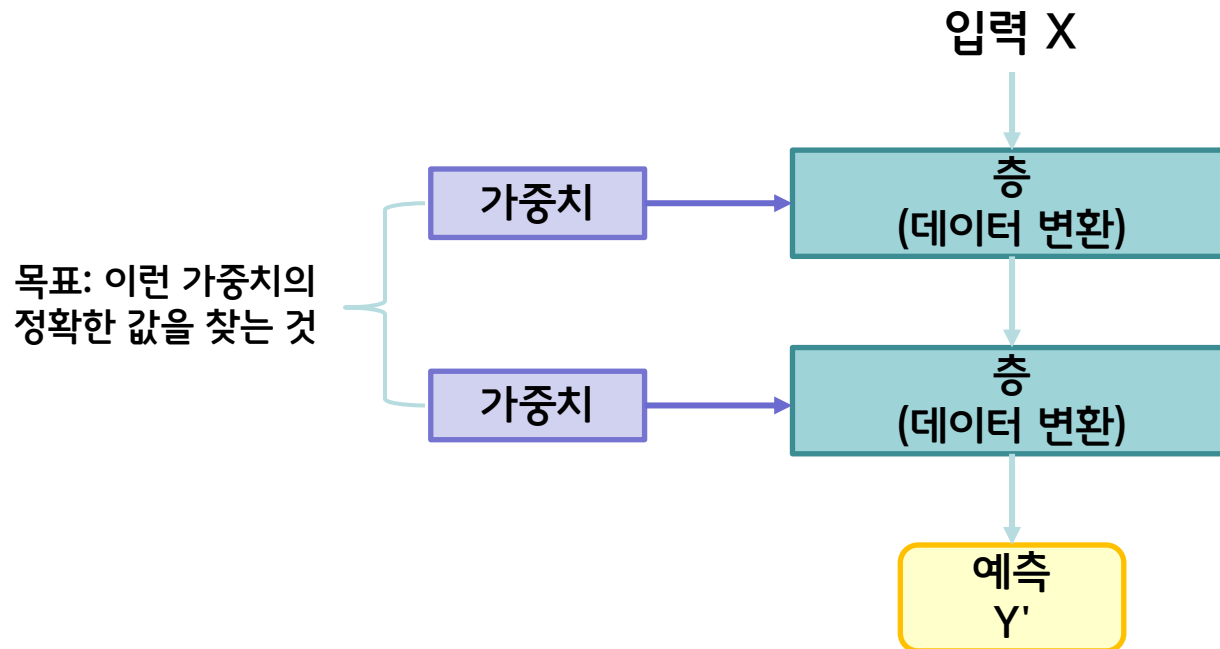
Multi-layer perceptron (MLP)는 Hidden layer라는 layer를 도입해 인풋을 한 차원 높은 단계의 특징, 즉 representation으로 나타낸다.



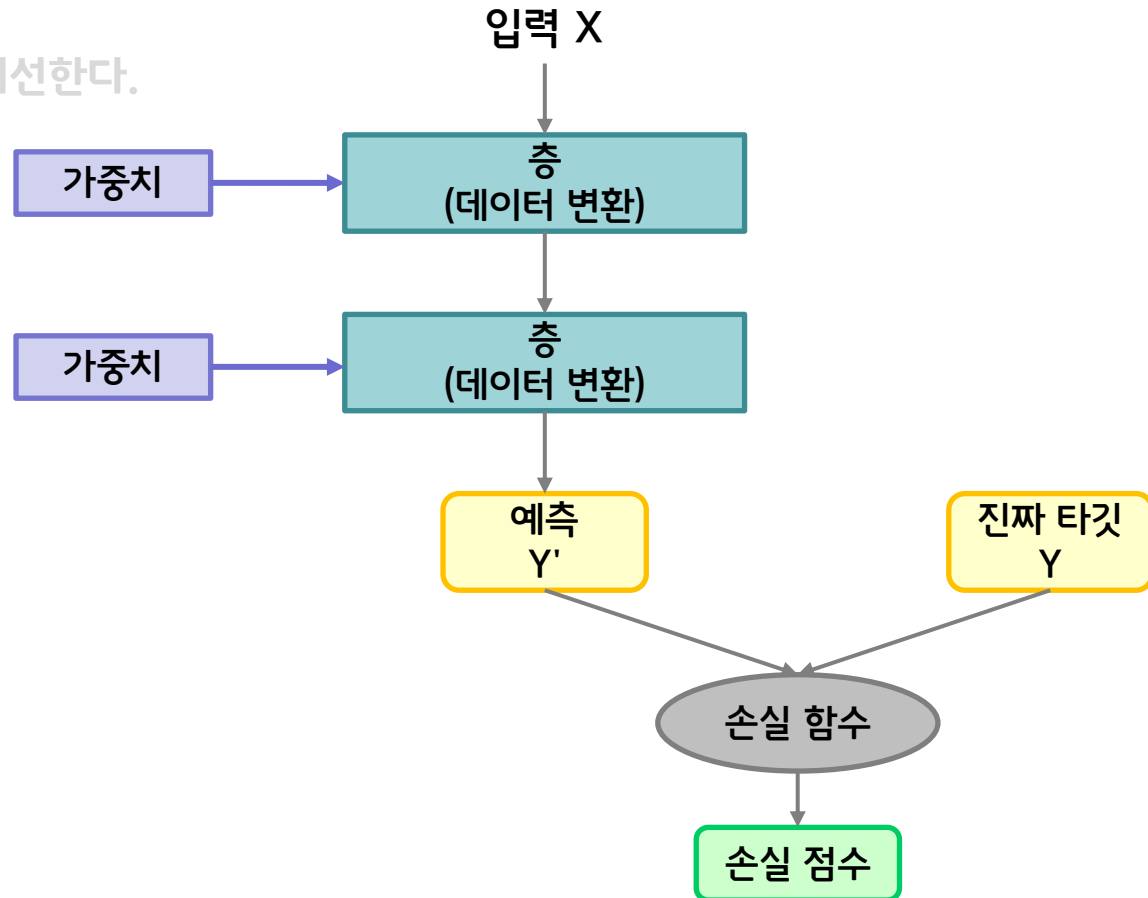
- ◆ **입력 데이터 포인트** : 음성인식에 있어서는 녹음된 사운드 파일, 이미지 태깅 작업에는 사진을 의미 한다.
- ◆ **기대 출력** : 음성 인식 작업에서는 사람이 사운드 파일을 듣고 옮긴 글, 이미지 작업에서 기대하는 출력은 '강아지','고양이' 같은 태그
- ◆ **알고리즘 성능 측정 방법** : 알고리즘의 현재 출력과 기대 출력 간의 차이를 결정하기 위해 필요하다. 피드백을 위해 필요한 것으로 이런 수정 단계를 학습(learning)이라고 한다.

- ◆ **타깃(target)** : 기대 출력을 의미한다.
- ◆ **매핑(mapping)** : 입력과 타깃의 관계로 입력을 representation로 변환, 연관시키는 것을 의미한다.
- ◆ **가중치(weight)** : 머신 러닝, 딥러닝 모두 결국은 가장 효율적인 식을 찾는 것이 목표이며, 이런 식 또는 식에 필요한 파라미터를 칭한다.
- ◆ **손실함수(loss function)** : 타깃과 출력 값의 차이를 계산하는 함수이다
- ◆ **역전파(Backpropagation)** : 손실함수의 결과를 개선하기 위해서 다시 결과에서부터 가중치를 수정하는 과정이다. 이를 옵티마이저(optimizer)가 담당한다.

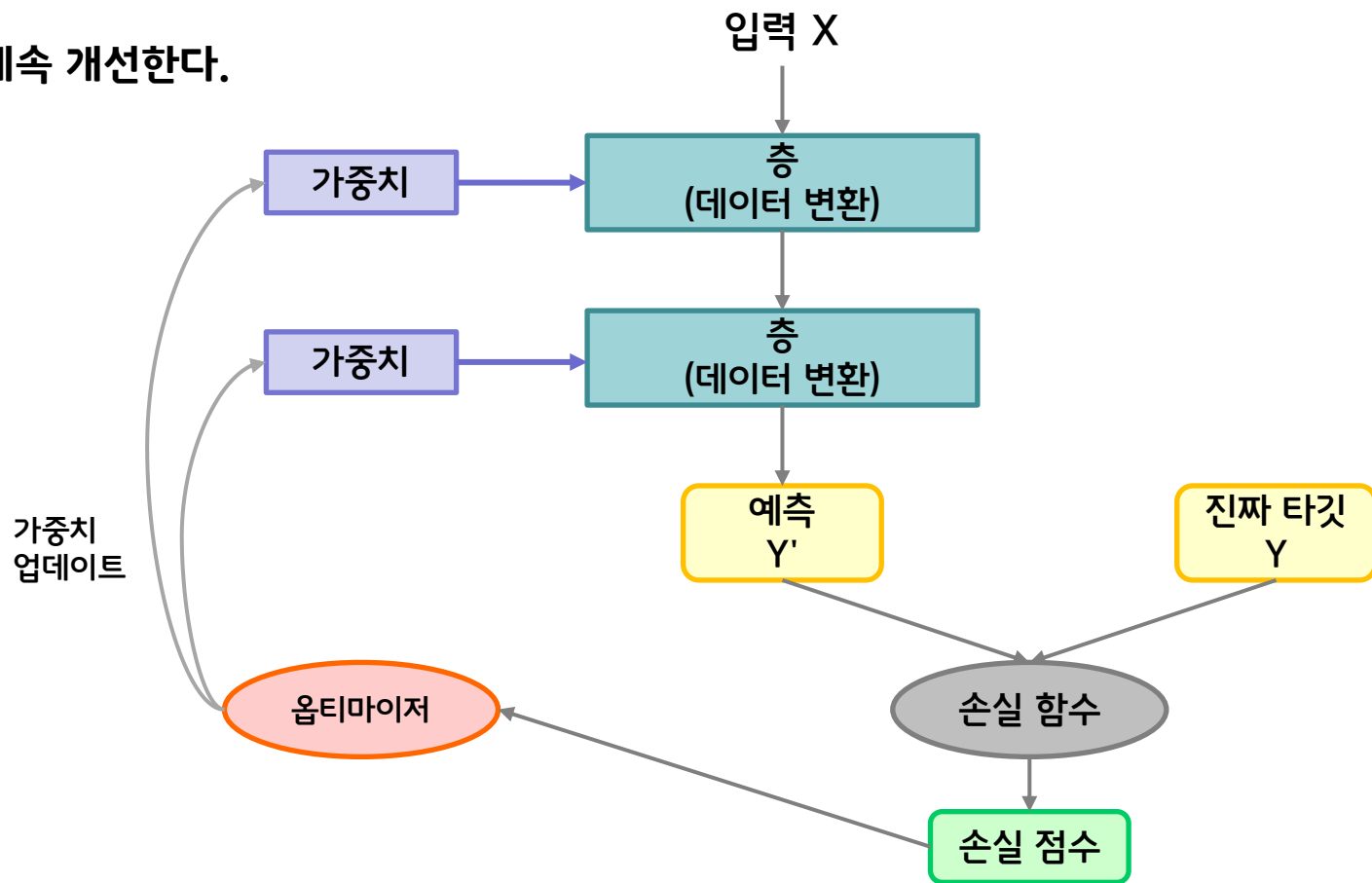
- 1.데이터를 입력한다.
- 2.여러 층을 통해 예상 결과값을 만든다. (매핑)
- 3.실제 값과 비교해서 그 차이를 구한다. (타겟과 손실함수)
- 4.차이를 줄이기 위한 방법으로 앞의 층들의 가중치를 수정해준다. (역전파)
- 5.이 방법의 반복으로 규칙을 계속 개선한다.



- 1.데이터를 입력한다.
- 2.여러 층을 통해 예상 결과값을 만든다. (매핑)
- 3.실제 값과 비교해서 그 차이를 구한다. (타겟과 손실함수)
- 4.차이를 줄이기 위한 방법으로 앞의 층들의 가중치를 수정해준다. (역전파)
- 5.이 방법의 반복으로 규칙을 계속 개선한다.



- 1.데이터를 입력한다.
- 2.여러 층을 통해 예상 결과값을 만든다. (매핑)
- 3.실제 값과 비교해서 그 차이를 구한다. (타겟과 손실함수)
- 4.차이를 줄이기 위한 방법으로 앞의 층들의 가중치를 수정해준다. (역전파)
- 5.이 방법의 반복으로 규칙을 계속 개선한다.



- ◆ 사람 수준의 이미지 분류, 음성 인식, 필기 인식
- ◆ 향상된 번역
- ◆ 향상된 TTS 변환
- ◆ 디지털 비서
- ◆ 자율 주행 능력
- ◆ 광고 타게팅
- ◆ 웹 엔진 결과
- ◆ 자연어 질의 대답 능력
- ◆ 바둑

- ◆ 지나친 기대는 큰 실망을 가져온다.
- ◆ 실망은 투자 감소로 이어진다.
- ◆ 투자 감소는 AI 겨울로 이어진다.

기술에 대한 거품이 증가하여, 갑작스럽게 지원이 많아지다 단기간내 성과가 없으면 혹 모두 투자를 안 하는 상황이 올 수 있다는 것이다.

이미 2번의 AI 겨울을 겪었고, 현재 3번째 겨울이 진행이 되고 있을지도 모른다는 점이다.

단기간의 기대는 비현실적이지만 장기적인 전망은 매우 밝다.

1. 딥러닝이란 무엇인가?

1.1 인공지능과 머신러닝, 딥러닝

1.2 딥러닝 이전 : 머신 러닝의 간략한 역사

1.3 왜 딥러닝일까? 왜 지금일까?

◆ 확률적 모델링(probability modeling)

- 통계학 이론을 데이터 분석에 응용한 것
- 초창기 머신 러닝 형태 중 하나이고, 현재에도 많이 사용됨
- 가장 잘 알려진 알고리즘은 나이브 베이즈(Naive Bayes) 알고리즘

◆ 로지스틱 회귀(logistic regression)

- 현대 머신 러닝의 "hello world"
- 이름은 회귀인데 회귀(regression) 알고리즘이 아닌 분류(classification) 알고리즘임
- 데이터 과학자가 분류 작업에 대한 초기 감을 위해 첫 번째로 선택되는 알고리즘임

나이브 베이즈 알고리즘이란 입력 데이터의 특성이 모두 독립적이라고 가정하고 베이즈 정리(Bayes' theorem)를 적용하는 머신 러닝 분류 알고리즘이다.

◆ 1950년대

- 신경망의 핵심 아이디어 등장
- 본격적으로 시작되지 못함

◆ 1980년대

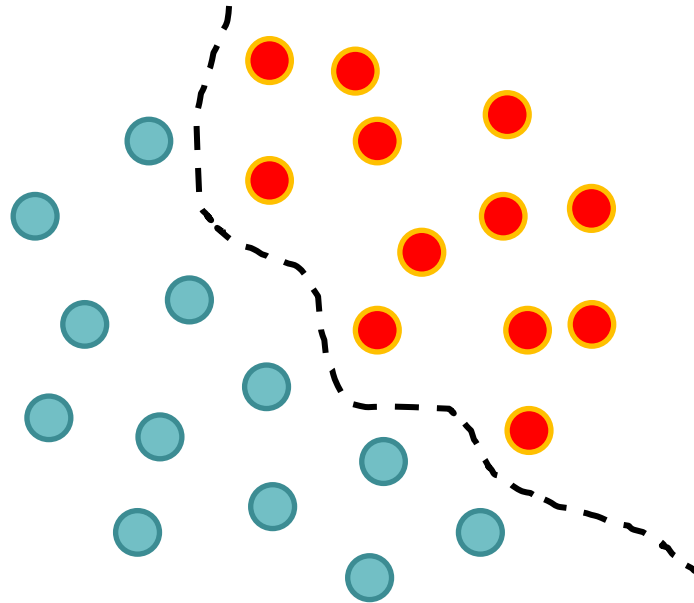
- 역전파 알고리즘 재발견
- 신경망에 역전파 알고리즘 적용 시작

◆ 1990년대

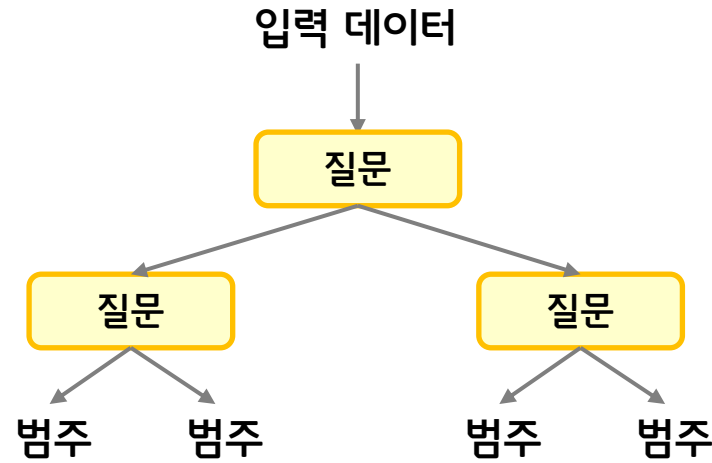
- 초창기 합성곱 신경망과 역전파를 연결
- 미국 우편 서비스에 이용

커널 방법(Kernel method) 분류 알고리즘의 한 종류로 그 중에서는 서포트 벡터 머신(Support Vector Machine, SVM) 이 가장 유명하다.

SVM은 분류 문제를 해결하기 위해 2개의 다른 범주에 속한 데이터 포인트 그룹 사이에 좋은 결정 경계(decision boundary)를 찾음



결정 트리(decision tree)는 플로우차트와 같은 구조를 가진다. 특히 랜덤 포레스트(Random Forest) 알고리즘은 결정 트리 학습에 기초한 것으로 안정적이고 실전에 유용하다.



그래디언트 부스팅 머신(gradient boosting machine)은 후에 나온 기법으로, 이전 모델에서 놓친 데이터 포인트를 보완하는 새로운 모델을 반복적으로 훈련함으로써 머신 러닝 모델을 향상하는 방법인 그래디언트 부스팅을 사용한다.

1. 딥러닝이란 무엇인가?

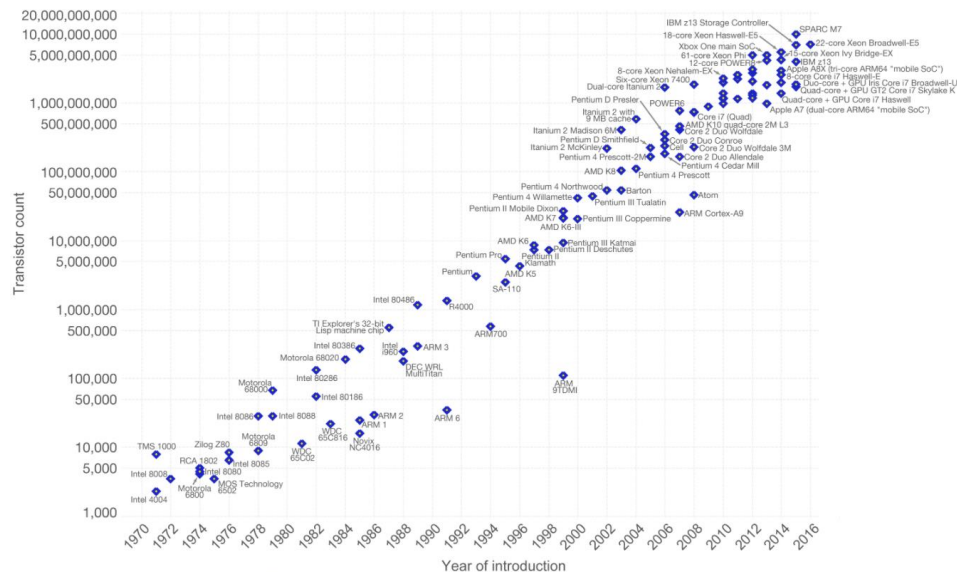
1.1 인공지능과 머신러닝, 딥러닝

1.2 딥러닝 이전 : 머신 러닝의 간략한 역사

1.3 왜 딥러닝일까? 왜 지금일까?

Moore's Law – The number of transistors on integrated circuit chips (1971-2016)

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore's law.



Data source: Wikipedia (https://en.wikipedia.org/wiki/Transistor_count)

The data visualization is available at OurWorldinData.org. There you find more visualizations and research on this topic.

Licensed under CC-BY-SA by the author Max Roser.

인텔의 고든 무어가 1965년에 주장한 법칙.

"반도체의 집적회로 성능은 18개월마다 2배로 증가한다"는 법칙이다.

이는 컴퓨팅 환경의 발전이 신경망의 이론을 뒷받침 하면서 이론으로만 존재 했던 여러 가지를 가능하게 만들었다.

CPU는 1990년부터 2010년 사이에 약 5000배 정도 빨라졌다.

2000년대 게임 그래픽 성능 개발을 위한 대용량 고속 병렬 칩(그래픽 처리장치 GPU)가 발전하였다.

GPU 제품을 위한 프로그래밍 인터페이스 CUDA를 출시하였다.

물리 모델링을 시작으로 신경망까지 병렬화가 가능해 졌다.

GPU인 NVIDIA TITAN X는 6.6 테라플롭의 단정도 연산 성능을 제공한다.

구글은 2016년에 텐서 처리 장치 프로젝트를 공개했다. 이 칩은 심층 신경망을 실행하기 위해 완전히 새롭게 설계한 것으로 최고 성능을 가진 GPU보다 10배 이상 빠르고 에너지 소비도 더 효율적이다.

2017에 발표한 TPU는 180 테라플롭이다.

‘데이터의 바다’라는 용어가 있듯이 현재는 데이터가 매우 많다.

저장 장치의 발전, 데이터 셋을 수집하고 배포할 수 있는 인터넷의 성장은 머신러닝에 필요한 데이터들을 마련할 수 있는 환경을 만들어주었다.

- 플리커에서 사용자가 붙인 이미지 태그
- 유튜브의 비디오
- 위키피디아는 자연어 처리 분야에 필요한 핵심 데이터셋
- 1400만개 이미지를 1000개의 범주로 구분해 놓은 ImageNet 데이터셋

- 신경망의 층에 더 잘 맞는 활성화 함수(activation function)
- 층별 사전 훈련(pretraining)을 불필요하게 만든 가중치 초기화
- RMSProp과 Adam 같은 더 좋은 최적화 방법 개발
- 배치 정규화
- 잔차 연결
- 깊이별 분리 합성곱

같은 고급 기술들이 개발 됨

초창기에 딥러닝을 하려면 흔치 않은 C++와 CUDA의 전문가가 되어야 했음

씨아노와 텐서 플로가 개발되어 JAVA나 Python으로 쉽게 개발할 수 있게 됨

케라스의 도구로 레고 블럭을 만들 듯 쉽게 새로운 모델을 개발할 수 있게 됨

케라스 등장(2015)이후 많은 스타트업과 학생, 연구자들이 활용함



단순함 : 딥러닝은 특성 공학이 필요하지 않아 복잡하고 불안정한 많은 엔지니어링 과정을 엔드-투-엔드로 훈련시킬 수 있는 모델로 바꾸어 준다.

확장성 : 딥러닝은 GPU 또는 TPU 에서 쉽게 병렬화할 수 있기 때문에 무어의 법칙 혜택을 크게 볼 수 있다. 또한 딥러닝 모델은 작은 배치 데이터에서 반복적으로 훈련되기 때문에 어떤 크기의 데이터셋에서도 훈련될 수 있다.

다용도와 재사용성 : 이전에 많은 머신 러닝 방법과는 다르게 딥러닝 모델은 처음부터 다시 시작하지 않고 추가되는 데이터로도 훈련할 수 있다.